METEOROLOGICAL OFFICE
- 5 JUN 1984
LIBRARY

LONDON, METEOROLOGICAL OFFICE.

Met.O.3 Technical Note No.28.

An evaluation of a simulation model of wind speed. By PARRINGTON,D.

London,Met.Off.,Met.O.3 Tech.Note No.28, 1984,31cm.Pp.19.9 Refs.Abs.p.3.

An unofficial document - restriction on first page to be observed.

ARCHIVE Y42.K2
National Meteorological Library and Archive
Archive copy - reference only

JB/10

MET O 3 TECHNICAL NOTE NO. 28

AN EVALUATION OF A SIMULATION MODEL OF WIND SPEED

BY

D PARRINGTON

(This report has not been published. Permission to quote from it should be obtained from the Assistant Director of the Climatological Services Branch of the Meteorological Office).

Climatological Services (Met O 3)

Meteorological Office

London Road

Bracknell

Berkshire

RG12 2SZ

March 1984

Contents

Summary

1. Introduction

2. The model

3. Persistence

4. Distribution of hourly mean windspeed

5. Extreme winds

6. Spells

7. Discussion

8. Conclusions

References

Summary

A model designed by Smith (1984) to simulate hourly mean wind speeds has been tested on various aspects of the wind climate corresponding to 3 main practical applications. The simulation of strong winds, for which the model was designed, is good but spells of light wind are poorly represented. The model also performs better at Elmdon, the site for which it was first developed, than for other locations. These findings illustrate the deterioration in the performance of the model as one passes from those aspects of the wind climate which were modelled explicitly to those that were not.

1. Introduction

The meteorological observations which form the data base for answering climatological enquiries are seldom homogeneous. Apart from direct changes in the site and instrumentation, a lack of homogeneity is caused by slow changes to the environment (e.g. to trees and buildings) and in the performance of instruments. Wind measurements are particularly vulnerable since they are very sensitive to their immediate environs and the level of instrumental maintenance. As a consequence, statistical models have been developed to generate large quantities of simulated wind data as the basis for answering climatological enquiries. Examples are the models constructed by Goh and Rathan (1979), Chou and Corotis (1981) and Smith (1984).

The stability of the input parameters in these simulation models is of paramount importance. Random errors in the input parameters will be converted to systematic errors in the generated data, and so a reasonably large number of observations are required in order to initialise the model. Usually it is only possible to model a small number of aspects of the real world explicitly, and errors will be increased whenever the use of the model is extended beyond these limits.

Strictly, a simulation model will be designed with set purposes in mind, and the use of the model for other purposes will be unjustified. Sometimes, however, the restrictions on the use of a model are unclear, and once the data have been simulated, there is the temptation to use them in a variety of contexts. The number of purposes to which simulated data may be legitimately applied is a very important property of a simulation model.

This paper examines the robustness of a model designed by Smith (1984) to generate mean hourly wind for any location in the UK. The model, which was not fully developed, was tested on 3 aspects of the wind climate corresponding to 3 main practical applications:-

i. The complete distribution of wind speed and its seasonal variation, which are of interest from the point of view of wind energy.

ii. Extreme winds, which are of importance in the design criteria of structures.

iii. Weather windows, i.e. periods of light winds, which are taken advantage of in the towing of fragile structures at sea.

Strong winds in general and extreme winds in particular correspond to aspects of the wind climate which were modelled explicitly, and which are therefore expected to be well simulated. Weather windows, on the other hand, correspond to a marine application for which the model was not designed, and may be poorly simulated.

The model was developed using data for Elmdon and subsequently generalised, so it is expected to perform better for Elmdon than for other sites. It is difficult to test the model on independent data since most of the suitable observations in the UK were used by Smith in the development of the model. Elmdon, Valley, Thorney Island and Aberporth were the stations chosen to evaluate the model. Tests at Elmdon provide an upper limit to the performance of the model, while a climatological contrast is provided by Valley, which was also used in the development of the model. Thorney Island and Aberporth represent independent stations, although the latter site has only 8 years of record available.

2. The model

Three distinct wind regimes were introduced to generate the sequences of mean hourly wind speed.

i. Depression events which represent the effects of synoptic features and which are associated with the higher wind speeds.

ii. A daily cycle in which the effects of solar heating on the wind is modelled by a truncated sine wave.

iii. Non-forcing conditions during which the wind speed is determined by a monthly average speed, the previous hours value, and a random component.

A very large number of input parameters are required to initialise the model, although they are not all independent:-

i. The cumulative probabilities of the amplitude of the daily cycle not exceeding a given value for various windspeeds at sunrise. The provision of 16 probabilities for 5 wind speed intervals is required, i.e. 80 numbers in all.

ii. The mean amplitude of the daily cycle in non-forcing conditions.

iii. The scale parameters of the 2 parameter gamma distributions used to model the peak wind speeds and separations of depression events.

iv. The monthly mean wind speed and its standard deviation.

v. The standard deviation of the random component and the first serial correlation coefficient in non-forcing conditions.

The items listed in i. to iv. are determined from raw data and involve the provision of 85 input parameters for each month. The parameters described in v. were arbitrarily set to 1.5 knots and 0.90 respectively.

It can be seen that the large number of input parameters requires a very considerable preliminary analysis of raw data. In order to obviate this, Smith provided a list of input parameters representing two zones which distinguished between western coastal and other districts. These were based on median values from 8 stations in each of the zones for items in iii. and iv., but were obtained only from Elmdon and Valley for items i. and ii. These parameters provide output appropriate to a location with the site and terrain characteristics of Elmdon. The simulated winds can then be adjusted to a site specific value using procedures described by Caton (1976) which involve factors such as altitude, gust ratio, and effective height of the anemometer. The zone parameters provided by Smith were used in this paper to generate the simulated winds.

3. Persistence

Conventional statistical tests assume that the data supplied are independent. Consecutive hourly mean wind speeds are, of course, highly correlated, and this serial correlation can be dealt with in two ways:-

i. By considering only observations which are so far apart that they are practically independent.

ii. By calculating the effective number of indepenent observations N_I in a series of N_R observations, where $N_I = N_R/F$.

The latter procedure is commonly used and is described in statistical texts such as Brooks and Carruthers (1953). For the simple Markovian case in which the nth serial correlation r_n is related to the first autocorrelation r through the equation $r_n = r^n$, it can be shown that

$$F = \frac{1+r}{1-r}$$

For the 4 stations examined, r = 0.945 giving F = 35.4.

The observed serial correlations at Elmdon are compared with those obtained from the simple Markov model based on r = 0.945 in Fig. 1. It can be seen that the observed correlations are higher, suggesting that F could be increased to about 50. In the statistical tests which follow, however, the value of 35.4 was retained. The effect is to produce an overestimate of N_I and hence a stringent test of the model.

The X^2 test is used for assessing the goodness of fit of the simulated to the observed distributions, but this is also affected by persistence. X^2 is defined as

$$X_I^2 = \sum_{i=1}^{n} \frac{(O - E_i)^2}{E_i}$$

where O_i and E_i are the observed and expected number of independent events in one of n classes. If we have correlated observations then the associated X^2_R will be based on F x N_I observations,

i.e.

$$X^2_R = \sum_{i=1}^{n} \frac{(F.O_i - F.E_i)^2}{F.Ei}$$

$$= F \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$= F.X^2_R$$

i.e. the value of X^2 obtained with correlated observations should be divided by F.

4. Distribution of hourly mean windspeed

The seasonal variation of the more commonly used statistics relating to the observed and simulated distributions of hourly mean wind are presented in Fig. 2 for Elmdon and Valley. The differences between the observed and simulated means were within the 95% confidence limits at all 4 stations examined. The standard deviation of the mean hourly wind is simulated reasonably well at Elmdon, but not at Valley where the pronounced seasonal variation is not reproduced.

A median test was also performed on the monthly distributions with the additive rule for X^2 being used to determine the overall significance for each station. In this case, although the majority of individual months passed a median test at the 10% significance level, the differences for the others were sufficiently large to make every station fail the overall combined median test at this level. The simulated medians may be close enough to those observed for most practical purposes, but the differences between them are statistically significant due to the large sample size.

The annual frequency distributions of hourly mean windspeed at Elmdon and Valley are presented in Fig. 3. There is close agreement for strong winds, but there are discrepancies for lighter winds, particularly at Valley where the results are similar to those for Thorney Island and Aberporth. Part of the discrepancy in light winds is due to instrumental unreliability at speeds less than 5 knots. A X^2 test was used to assess the goodness of fit of the annual distribution both including and excluding speeds less than 5 knots. In both cases, the differences were significant at the 0.1% level for all the stations examined.

In order to check on the persistence correction, the X^2 test was repeated using observations selected at intervals of 24 to 168 hours. The sampling intervals were kept as multiples of 24 hours in order to ensure that the same population was sampled, and the hour of observation was made 12 noon in order to minimise the number of occasions with wind speeds less than 5 knots. The results, presented in Table 1, show that X^2 uncorrected for persistence stabilises when the sampling interval is 96 hours or more.

These values of X^2 are similar to those obtained by dividing the unadjusted X^2 obtained from all observations by F = 35.4. This suggests that the correction procedure described in Section 3 is valid and that observations taken once every 4 days are practically independent.

5. Extreme winds

The highest mean hourly winds in each month and year were extracted from 100 years of simulated and 15 years of real data for 3 stations. The observations from Aberporth were not considered because of the brevity of the record. The monthly and annual maxima were ranked and a cumulative probability P was attached to the mth ranking of N observations from an application of the formula

$$p = \frac{m-0.31}{N+0.38}$$ (Jenkinson, 1969)

A conventional analysis of the extremes was then performed as described by Gumbel (1958). The ranked observations were plotted against a transformation of the probability p known as the reduced variate y given by

$$y = -\ln[\ln(p)]$$

A linear relation was fitted to the plot using a technique due to Lieblein (1974) and an example is shown in Fig. 4 using data for Valley. The simulated extremes were evaluated from a comparison of the location and scale parameters of the regressions derived from the observed and simulated

extremes, together with estimates of the wind speed to be expected once every 50 years. Volume 1 of the Flood Studies Report (NERC, 1973, p103) gives the standard errors of these parameters as

$$SE(U) = 1.05\alpha/\sqrt{N}$$

$$SE(\alpha) = 0.78\ \alpha/\sqrt{N}$$

and $$SE(T_{50}) = 3.52\alpha/\sqrt{N}$$

where U is the intercept, α is the slope and T_{50} is the event to be expected once every 50 years. As the observed and simulated extremes are uncorrelated, the standard errors of the differences between the parameters derived from the observed and simulated observations can be obtained from the addition of variances. The extreme value analyses of the observed and simulated observations are compared in Fig. 5 by expressing the differences of U, α and T_{50} in terms of these standard errors. It can be seen that the differences exceed 2 standardised anomalies on only 11 out of 108 occasions. This represents an acceptable simulation of extremes.

6. Spells

There are many possible ways of examining spells, but for the purpose of evaluating the model relatively simple methods were chosen. Comparisons of the observed and simulated records were made in respect of the average number and duration of spells in each month, followed by an examination of the complete distribution of the duration of spells. A spell is defined as

any continuous period during which the windspeed remains below a given threshold. A value of 10 knots was chosen for this because it did not produced spells which were either very long or very rare.

The average number and duration of spells starting in each month are presented in Fig. 6. This shows that the model produces too many short spells, which may be due to too strong a diurnal cycle and/or too many model days on which the diurnal cycle was important. The complete distribution of spell durations at Thorney Island and Aberporth is examined in Fig. 7, and this shows that too small a number of long spells are simulated. The result of X^2 tests to assess the goodness of fit of the distribution of spell durations are reported in Table 2. The best findings are for Aberporth, where in the majority of months the differences did not reach the 5% significance level. At Valley, however, all months recorded differences at the 0.1% level of significance. Overall, the results show that the model does not generate reliable statistics of spells of wind speed less than 10 knots.

Since strong winds in depression events are generally well simulated, the performance of the model in generating spells of wind exceeding large thresholds is likely to be better than that found above. Most practical interest in spells, however, relates to periods of light winds at sea. Since the failure of the spells analysis presented above may have been contributed to by too strong a diurnal cycle, the results in a maritime environment, where the daily cycle is largely absent, may be better. As a consequence, the daily cycle was removed from the model before it was used

to generate winds at Lerwick. The results, however, were similar to those presented above, i.e. there were too many short spells and too few long ones.

7. Discussion

It is important that a distinction be made between a practical and a statistically significant difference. Given a large enough sample, any difference will be statistically significant regardless of how small it is. On the other hand, with a small sample size, quite large differences may not be statistically significant. If large samples are being used in a statistical test the magnitude of the difference must be considered before a variable is rejected since a difference which is statistically significant may be trivial in practice. Whenever small samples are used in statistical testing it should always be remembered that a failure to produce a signficant difference may mean that there are insufficient data to determine whether the differences are due to chance. In the statistical testing of the complete distribution of windspeed, the large sample size renders statistically significant differences which are unimportant in practice, while for the extreme values the small sample size reduces the ability of the statistical tests to provide a useful assessment of the model.

A point of interest relating to extreme winds is that they are drawn from two populations. Most are associated with depression events but in summer several are produced by the diurnal cycle. This phenomenon, however, is not restricted to the model but occurs in nature as well.

A weakness of the model pointed out by Smith is the underestimate of the year to year variability, suggesting the need for the inclusion of a persistence term. The separations of depression events are based on observations taken from individual calendar months. In summer these events are rare with separations approaching a month. A better simulation of these separations might therefore be made by basing the distribution on observations made over a longer period of time than one month.

The amplitude of the diurnal cycle in the model is controlled by the month and the windspeed at sunrise. A table of cumulative probabilities defines the likelihood of the amplitude being a particular value for various intervals of sunrise speed. The two intervals that have the greatest effect on spells of windspeed below 10 knots, and low windspeeds in general, are sunrise windspeeds below 3.5 knots and those between 3.5 knots and 7.5 knots. Since the observed windspeed record is unreliable below 5 knots due to instrumental deficiencies, an improvement in the simulation might be made by moving the 3.5 knot class divider out of the questionable region to 5 knots.

Two of the input paramters required by the model are defined somewhat arbitrarily. These are the standard deviation of the random component and the first serial correlation coefficient in non forcing conditions. The first autocorrelation coefficient in all conditions at the stations examined ranged from 0.93 to 0.97 and the value under non-forcing conditions may be expected to be a little less than this. Consequently the value imposed, 0.90, is probably reasonable.

The standard deviation of the random component has been set to 1.5 knots. It would be better if it were made a function of the simulated wind speed as random variations are likely to be greater in strong rather than light winds. For most stations and seasons, this would result in an increase in the random component. A deficiency of the model is its inability to reproduce seasonal variations in the standard deviation of hourly mean wind speed. The standard deviation of the simulated winds is largely determined by the random component, and the indicated change to the latter would enable the standard deviation to vary from month to month.

The random number generator used in producing the random component gives a normally distributed set of numbers. It is possible that a different choice of random number generator, e.g. one producing a skewed distribution, would lead to improved simulations. The provision of new parameters for more clearly defined districts (zones) of the UK would probably also lead to improved estimates of wind speed.

8. Conclusions

The results clearly demonstrate how the performance of the model deteriorates as we pass from those aspects of the wind climate which were modelled explicitly to those that were not. Smith's main practical aims were concerned with wind energy and design criteria. Although the demands of wind generators depend on the overall distribution of wind speed, they are far more sensitive to periods of strong rather than light winds.

Consequently Smith directed most of his efforts towards modelling strong winds. He points out that it is very difficult to model light winds due to their dependence on topography and instrumental limitations.

From the above considerations, it is not surprising that extreme winds constitute the feature of the wind climate most succesfully simulated by the model. The upper portion of the complete distribution is also well modelled, but at lower speeds considerable differences from reality occur, and these result in the failure of the model to adequately simulate the complete distribution of wind speed. The other practical application examined relates to light winds at sea. This is an application for which the model was not designed and this is reflected in its poor performance in this respect.

The model is based on 3 types of wind regime - depression events, the daily cycle, and non-forcing conditions. The good simulation of strong winds is a result of the careful modelling of depression events. These only occur on about 5% of occasions, however, so for most of the time the simulated winds are generated by the daily cycle and random variations. An improved simulation of spells of light wind is only likely to be achieved by modelling them explicitly.

The model was developed first using data for Elmdon and was subsequently generalised to other sites. The model never reached the stage of full development and this is reflected in its better performance at Elmdon than

elsewhere. The design of a single model which will cope with the wide variety of wind climates found in the UK is, however, a very prodigious task.

References

Brooks, C.E.P. and Carruthers, N. 1953 Handbook of statistical methods in meteorology. HMSO, London.

Caton, P.F.G. 1976 Maps of hourly mean windspeed over the United Kingdom 1965–73. Clim. Mem. No. 79, Meteorological Office, Bracknell.

Chou, K.C. and Corotis, R.B. 1981 Simulation of hourly wind speed and array wind power. Solar Energy, 26, 199–212.

Goh, T.N. and Nathan, G.K. 1979 A statistical methodology for study of wind characteristics from a close array of stations. Wind Eng., 3, 197–206.

Gumbel, E.J. 1958 Statistics of extremes. Columbia Univ. Press, New York.

Jenkinson, A.F. 1969 Statistics of extremes. In estimation of maximum floods, WMO Tech. Note. No. 98, Chapter 5, 183–257.

Lieblein, J. 1974 Efficient methods of extreme value methodology. US Dept. of Commerce, National Bureau of Standards, Final Report NBSIR 74-602.

NERC 1975 Flood Studies Report, Vol. 1, London.

Smith, S.G. 1984 A statistical model to simulate sequences of hourly mean wind speeds over the United Kingdom. J. Clim., 4.

Table 1 - Goodness of fit of simulated to observed distribution of hourly mean windspeed for 1200 hours at Elmdon.

Sampling interval (hours)		24	48	72	96	120	144	168
χ^2 statistic	for all wind speeds	262	117	99	50	31	52	42
	excluding winds < 5 Knots	226	106	88	46	26	50	38

The χ^2 statistic was based on 6 degrees of freedom when using all observations and 4 degrees of freedom when excluding those less than 5 Knots. All the calculated values of χ^2 are significant at the 0·1% level.

Table 2 - Comparison of observed and simulated duration of spells with windspeed below 10 Knots.

	Elmdon		Valley		Thorney Island		Aberporth	
	χ^2	Sig(%)	χ^2	Sig(%)	χ^2	Sig(%)	χ^2	Sig(%)
January	12·3	3	22·3	<0·1	8·2	15	3·3	65
February	14·4	1	195·2	<0·1	13·7	2	4·8	44
March	26·3	<0·1	110·3	<0·1	24·7	<0·1	11·5	4
April	14·2	1	79·4	<0·1	28·8	<0·1	19·6	0·2
May	21·1	<0·1	99·8	<0·1	56·2	<0·1	10·1	7
June	18·8	0·2	135·9	<0·1	83·9	<0·1	26·5	<0·1
July	29·8	<0·1	65·8	<0·1	77·8	<0·1	13·9	2
August	6·5	26	62·3	<0·1	39·8	<0·1	9·8	8
September	6·8	23	109·9	<0·1	20·7	<0·1	11·0	5
October	16·6	0·5	127·9	<0·1	3·2	68	3·3	66
November	7·4	19	130·8	<0·1	5·7	34	8·8	12
December	6·9	23	95·5	<0·1	9·9	8	8·2	15

All χ^2 tests were based on 5 degrees of freedom.

Fig 1:- The Serial Correlation of Hourly Mean Wind Speed at Elmdon.

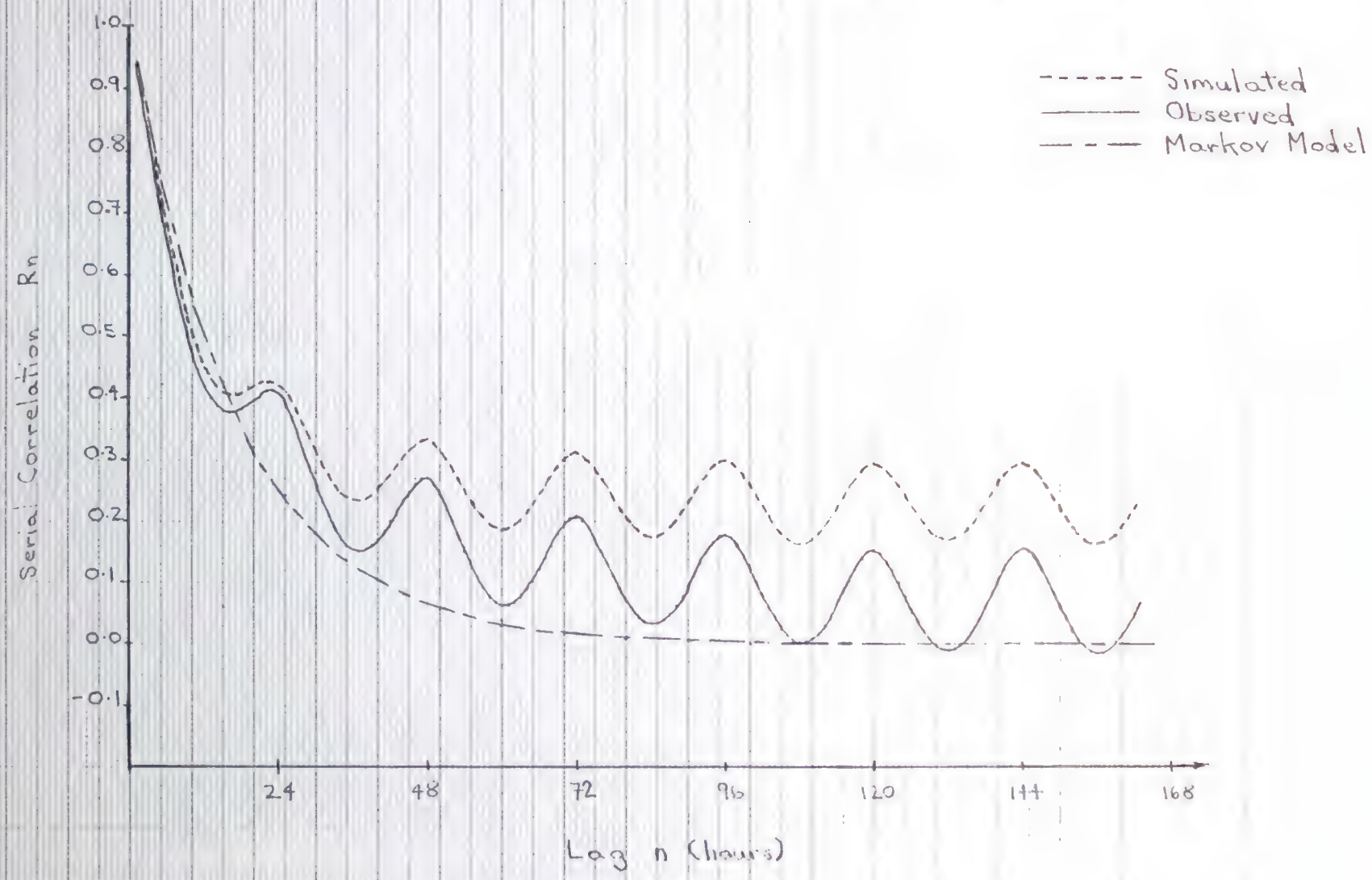

FIG 2 – CHARACTERISTICS OF HOURLY MEAN WIND AT ELMDON & VALLEY.

——— Observed - - - - Simulated

Error bars represent ± one standard error of observed values

ELMDON VALLEY.

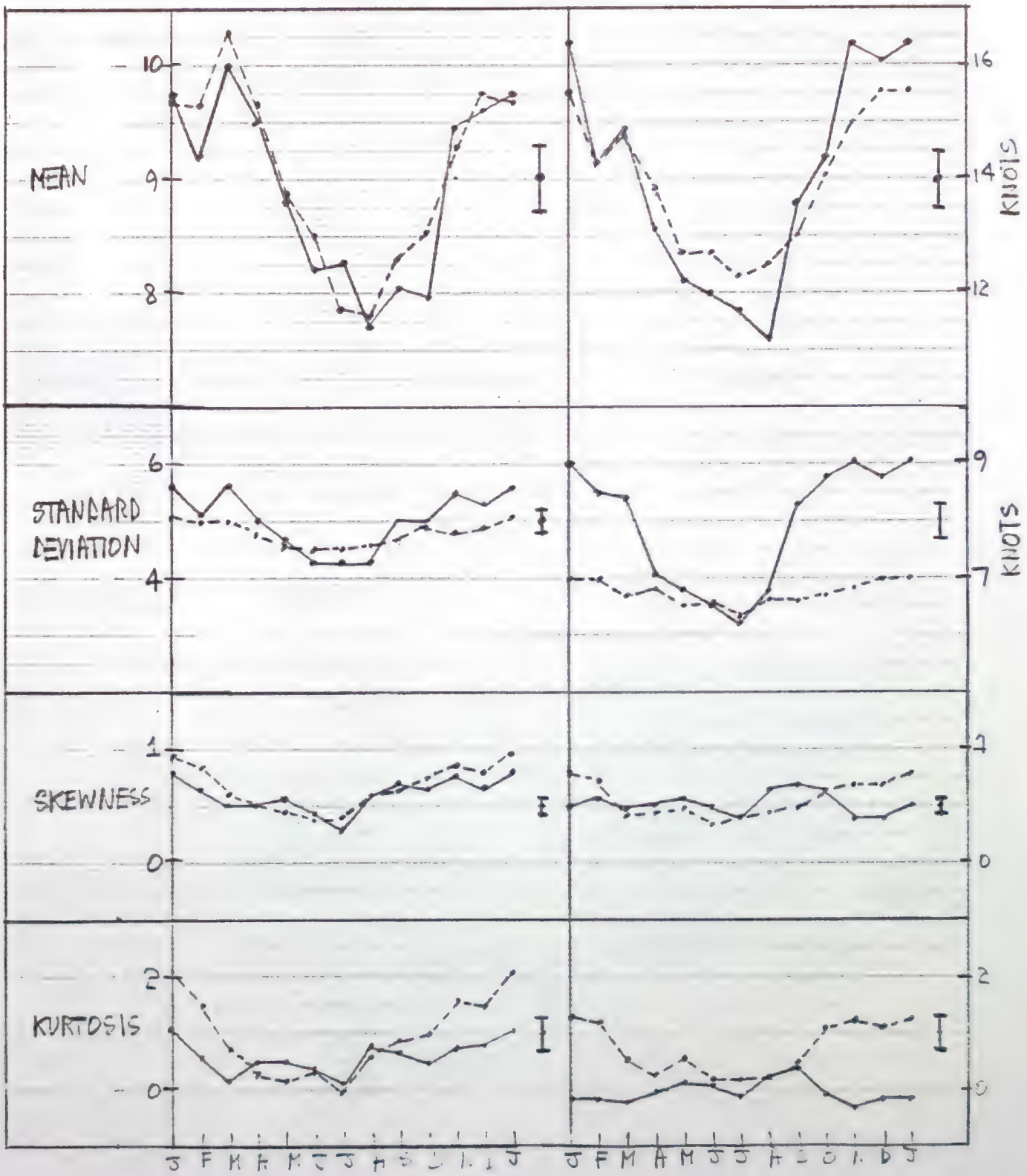

Fig 3:- The Distribution of Hourly Mean Wind Speed in a year

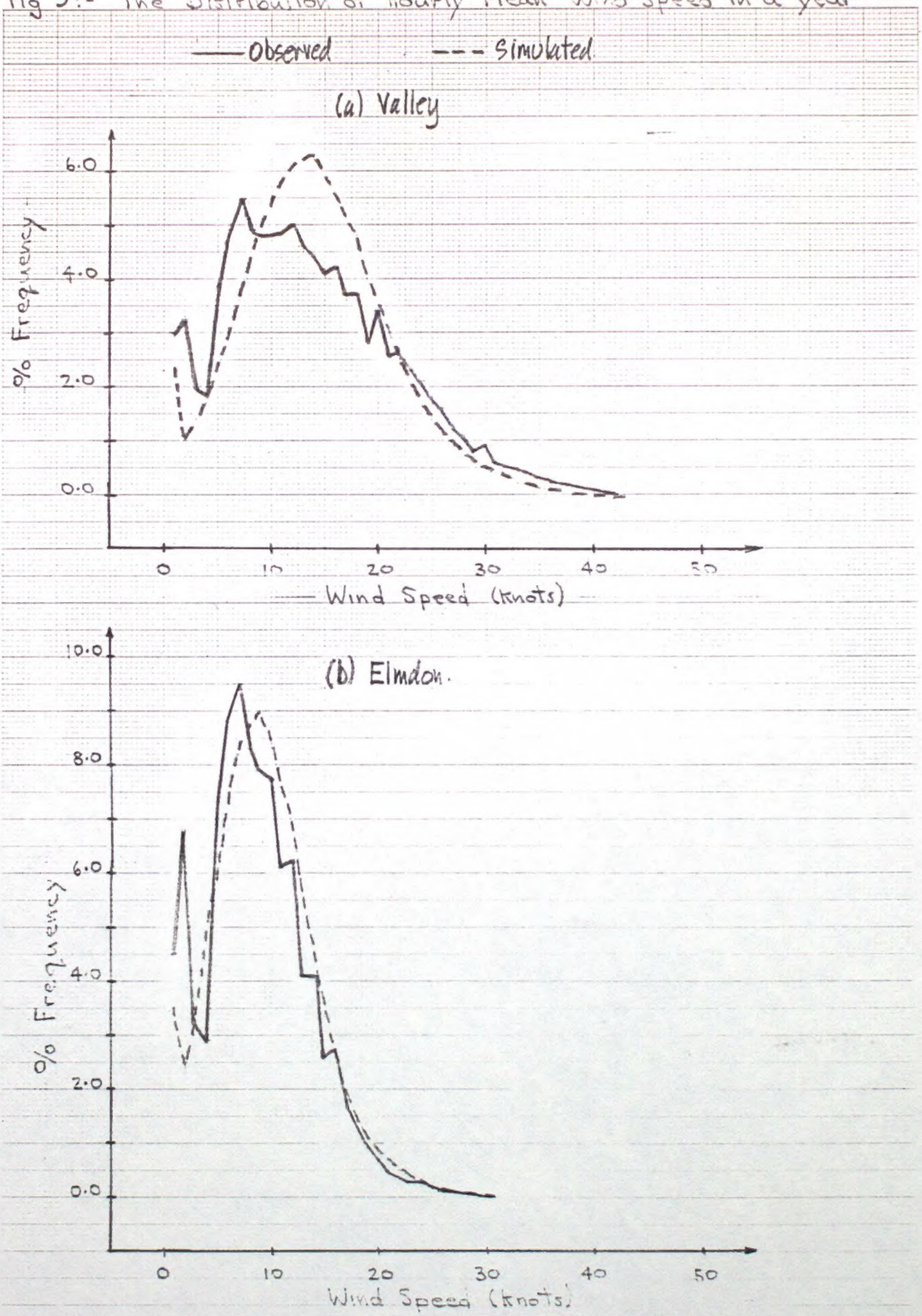

Fig 5:- Standardised Residuals of Extreme Wind Speed Parameters

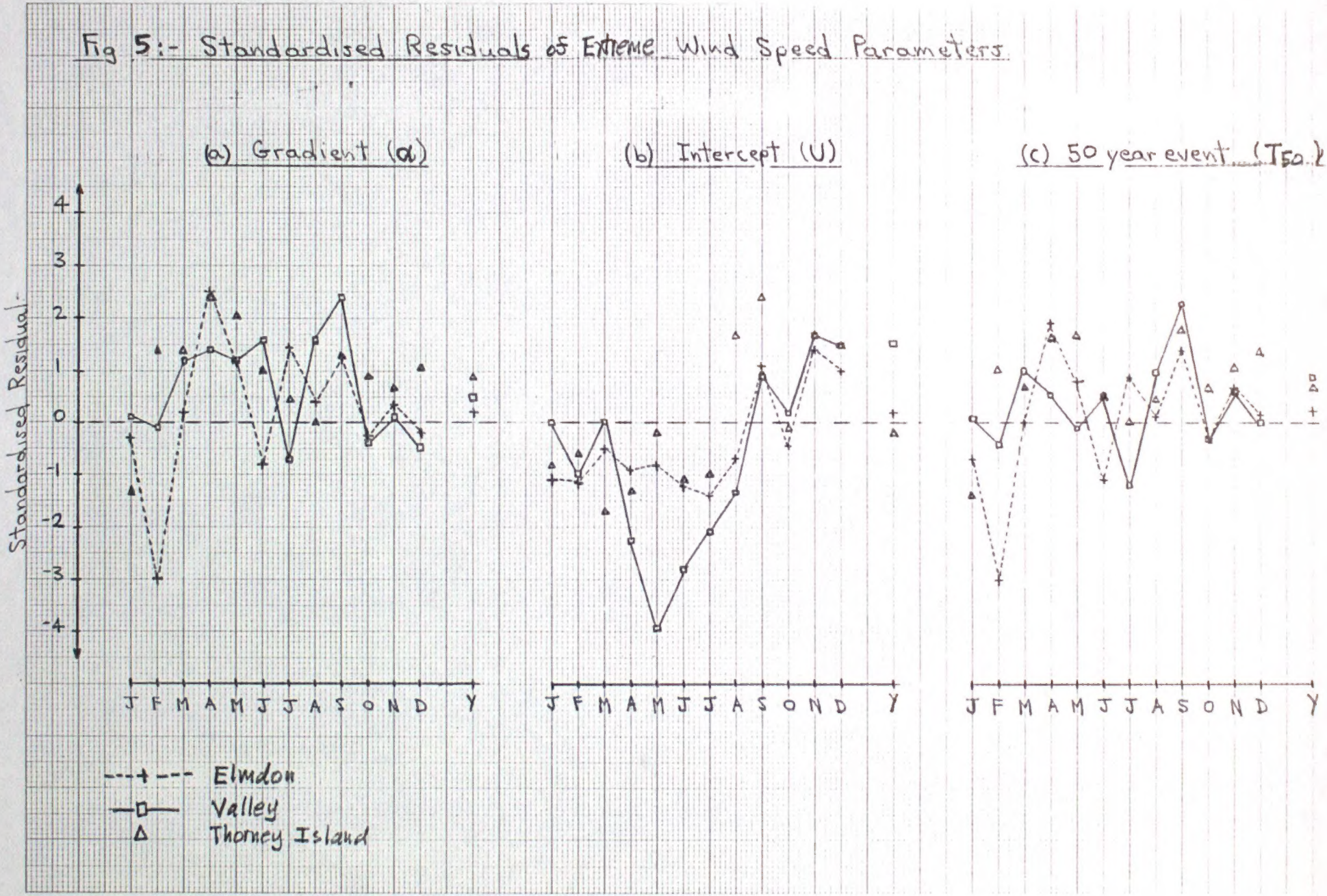

FIG 6 - FREQUENCY AND MEAN DURATION OF SPELLS WITH WINDSPEED LESS THAN 10 KNOTS

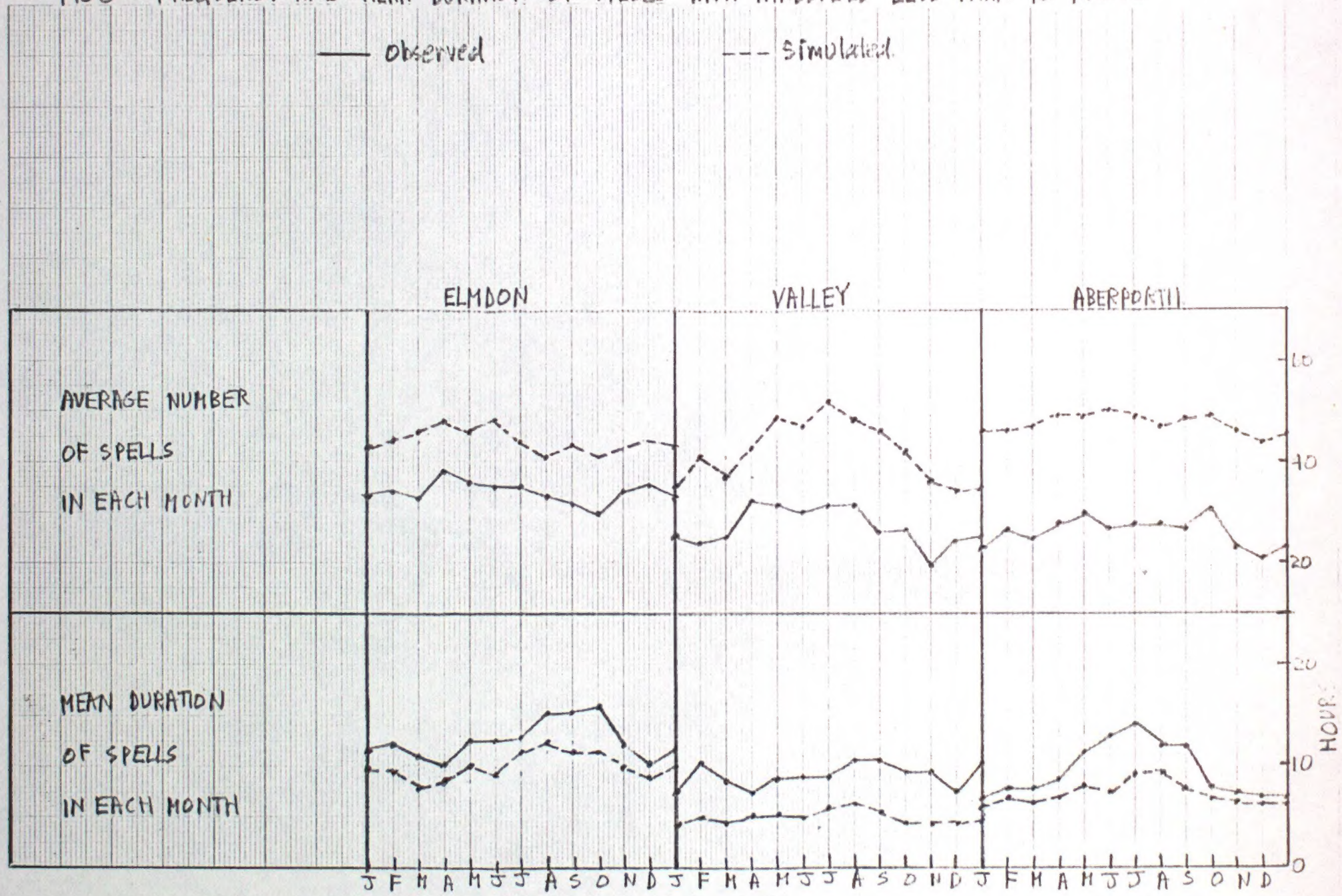
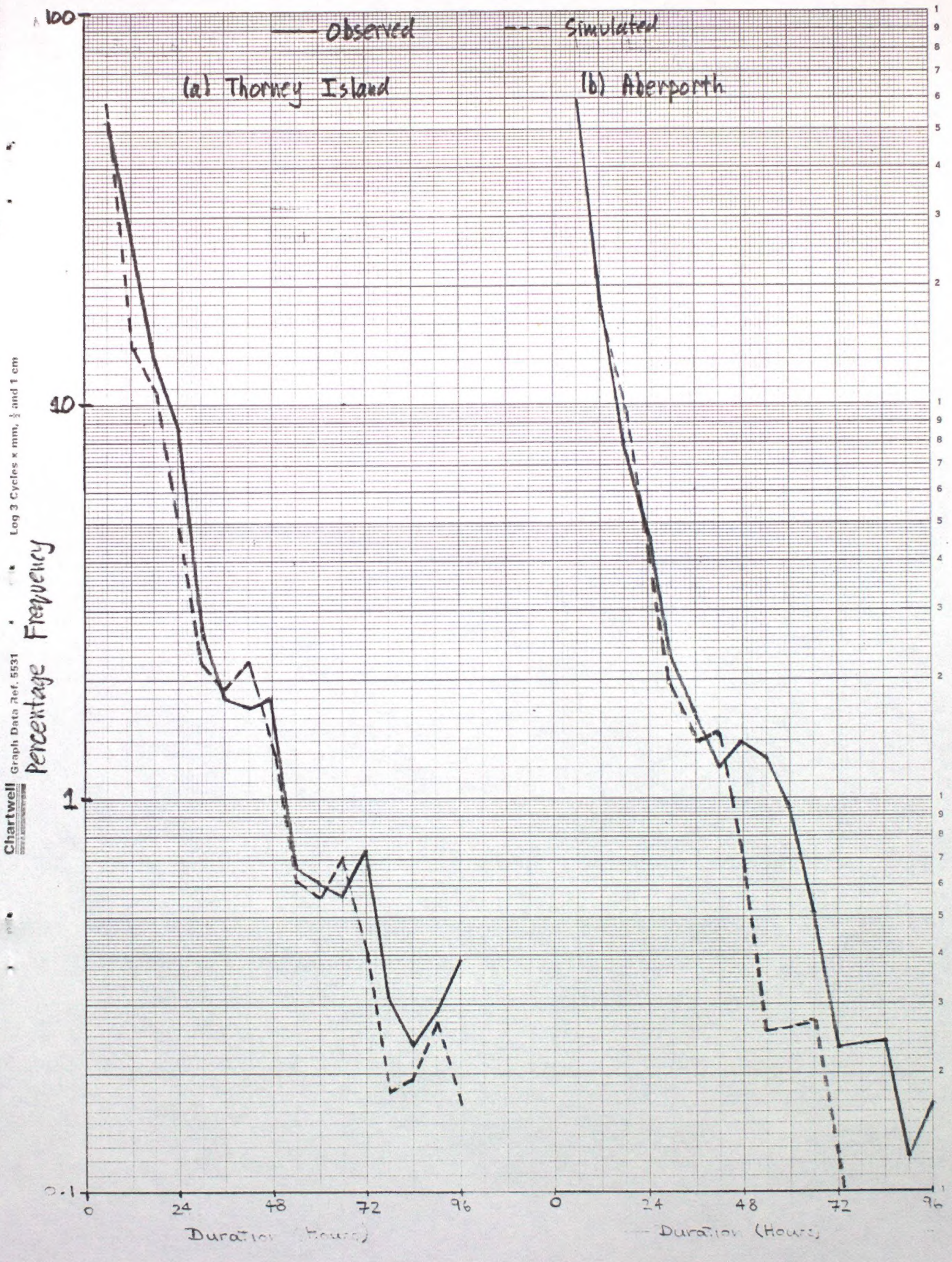
Fig 7 - The distribution of the duration of spells with windspeed less than 10 knots